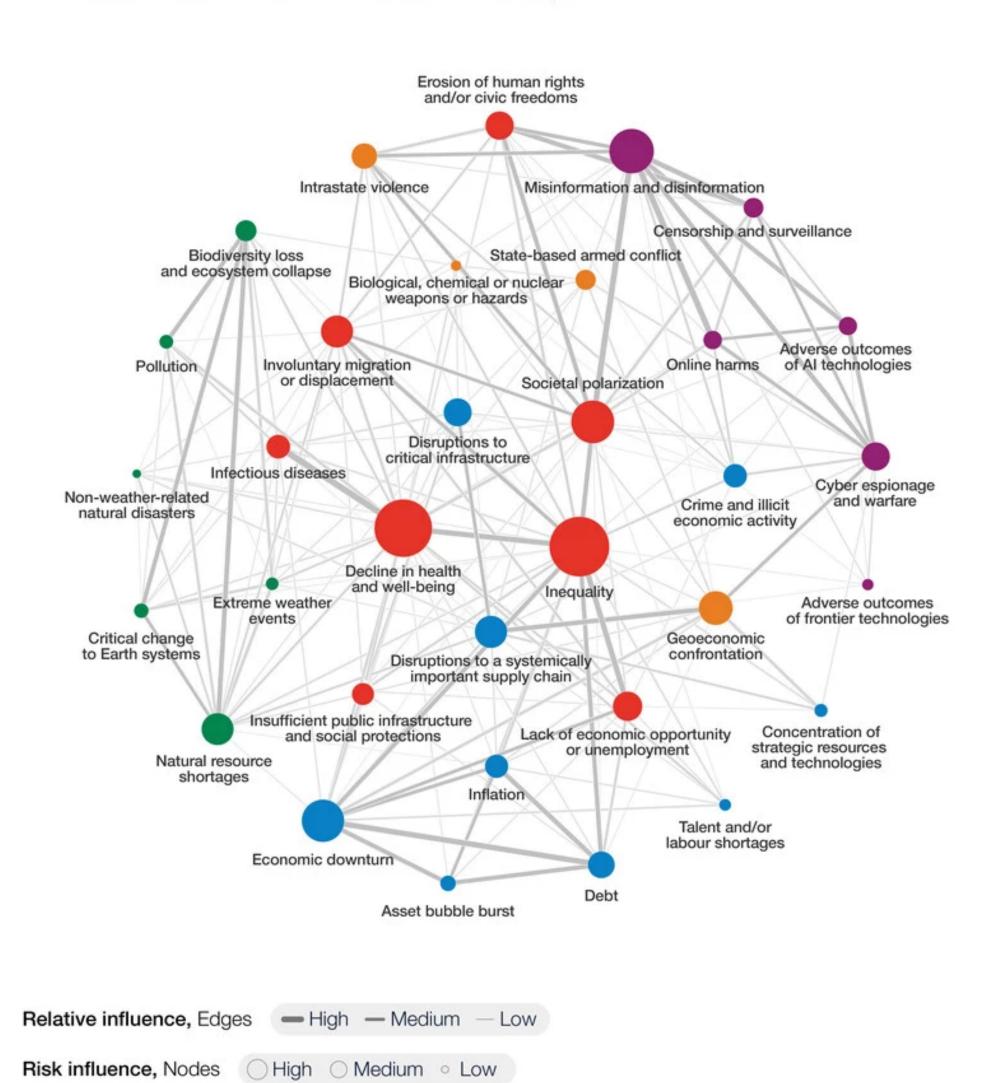
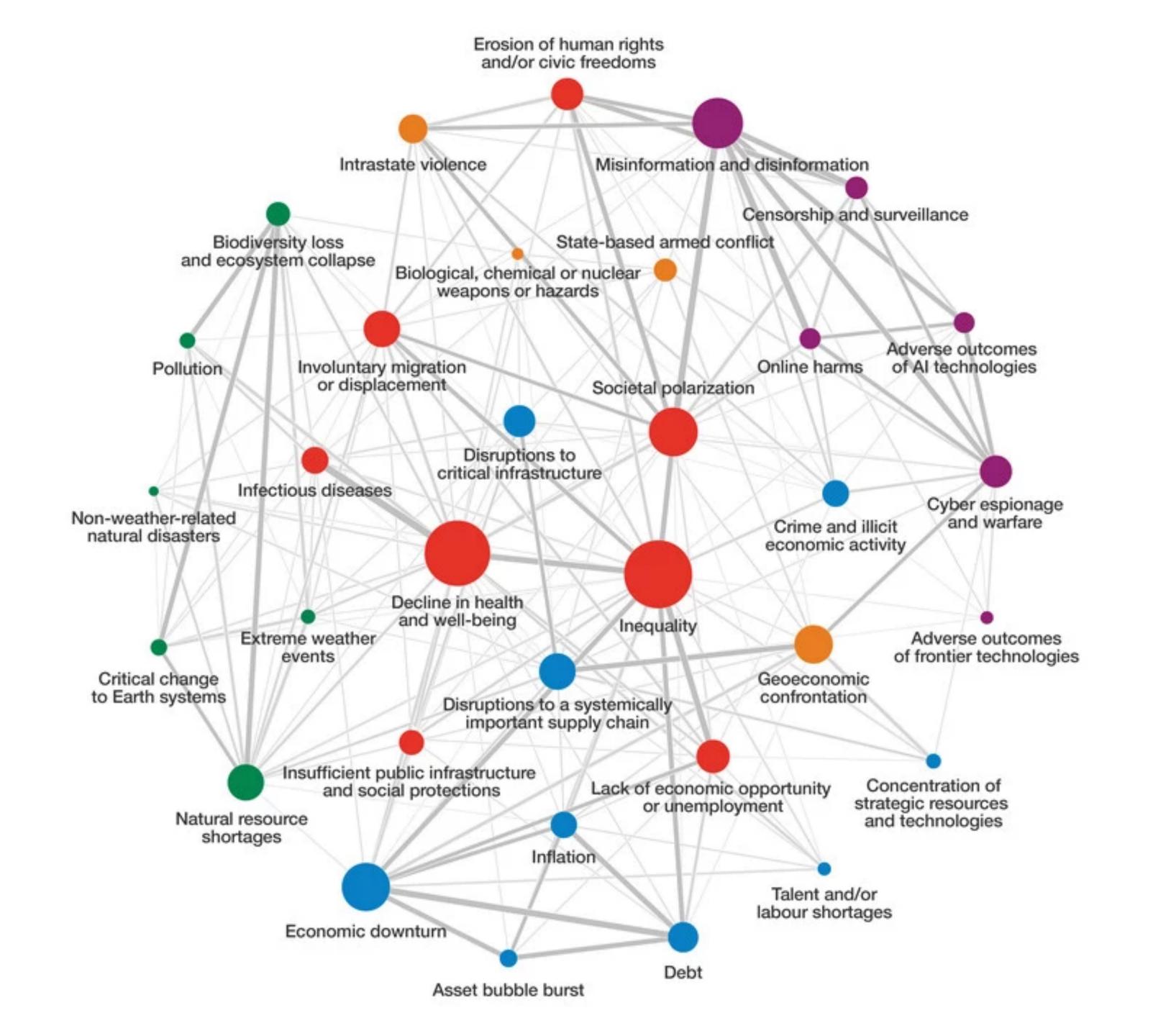
Action Research in the Age of GenAl

Ruben R. Puentedura, Ph.D.

Global risks landscape: an interconnections map

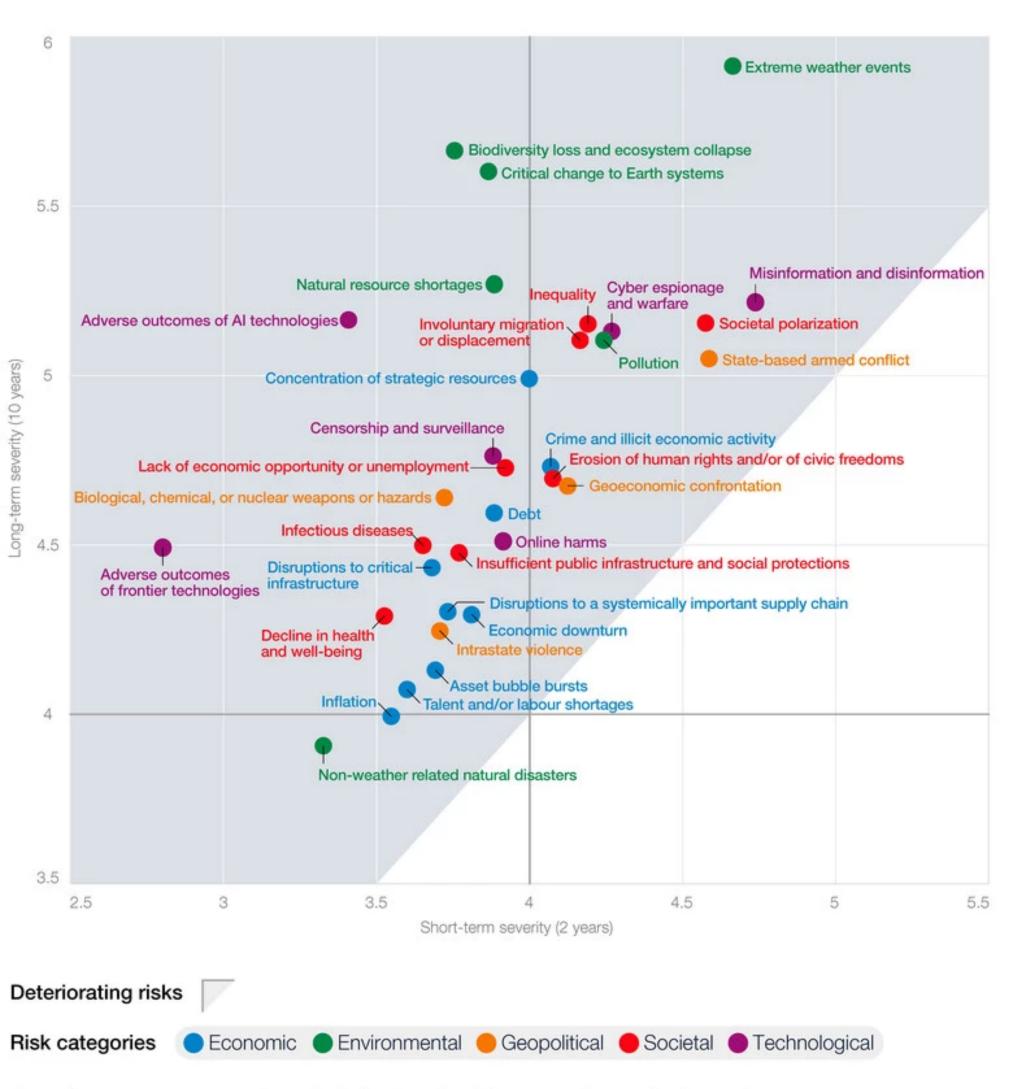






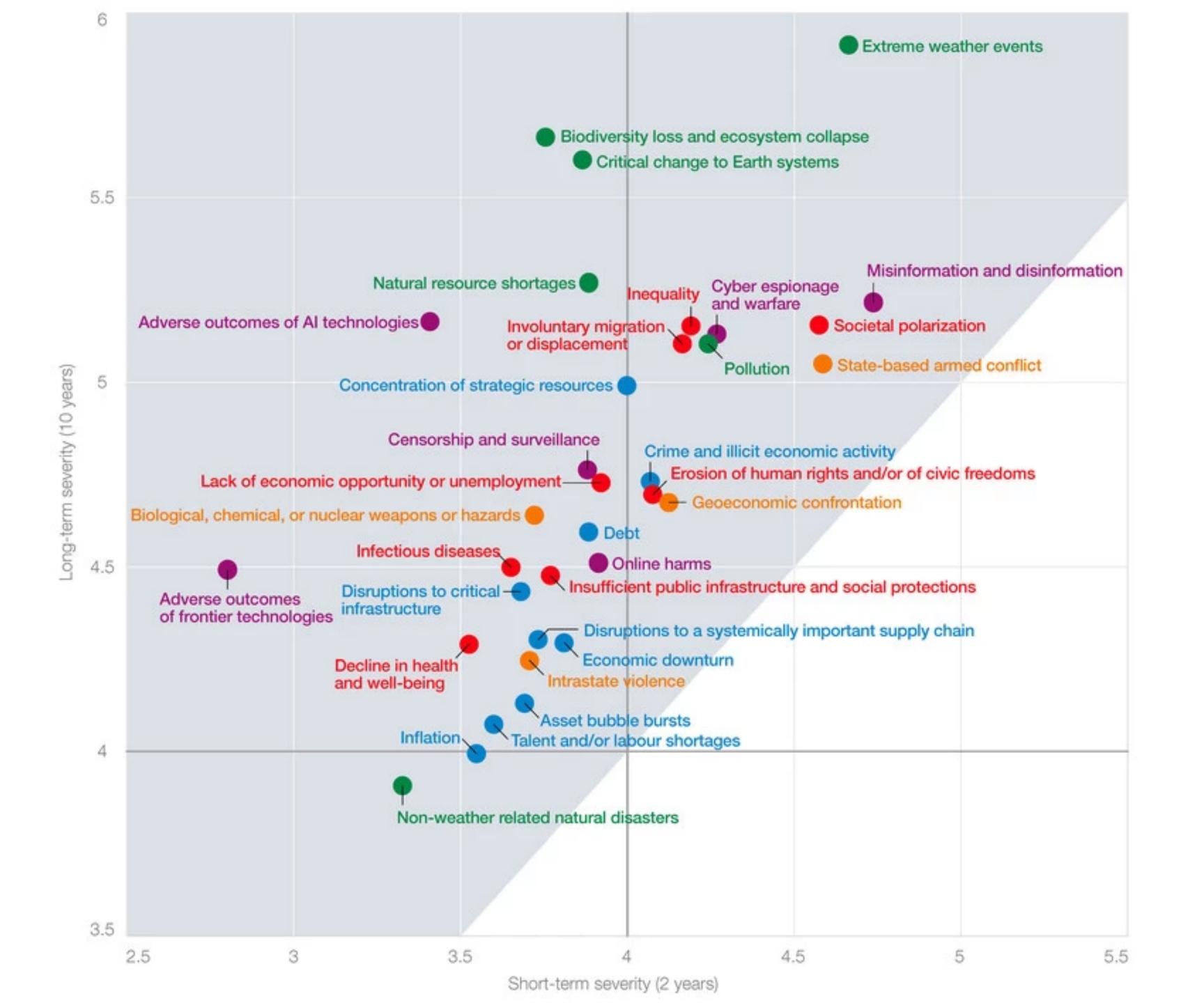
Risk Severity in 2 and 10 years

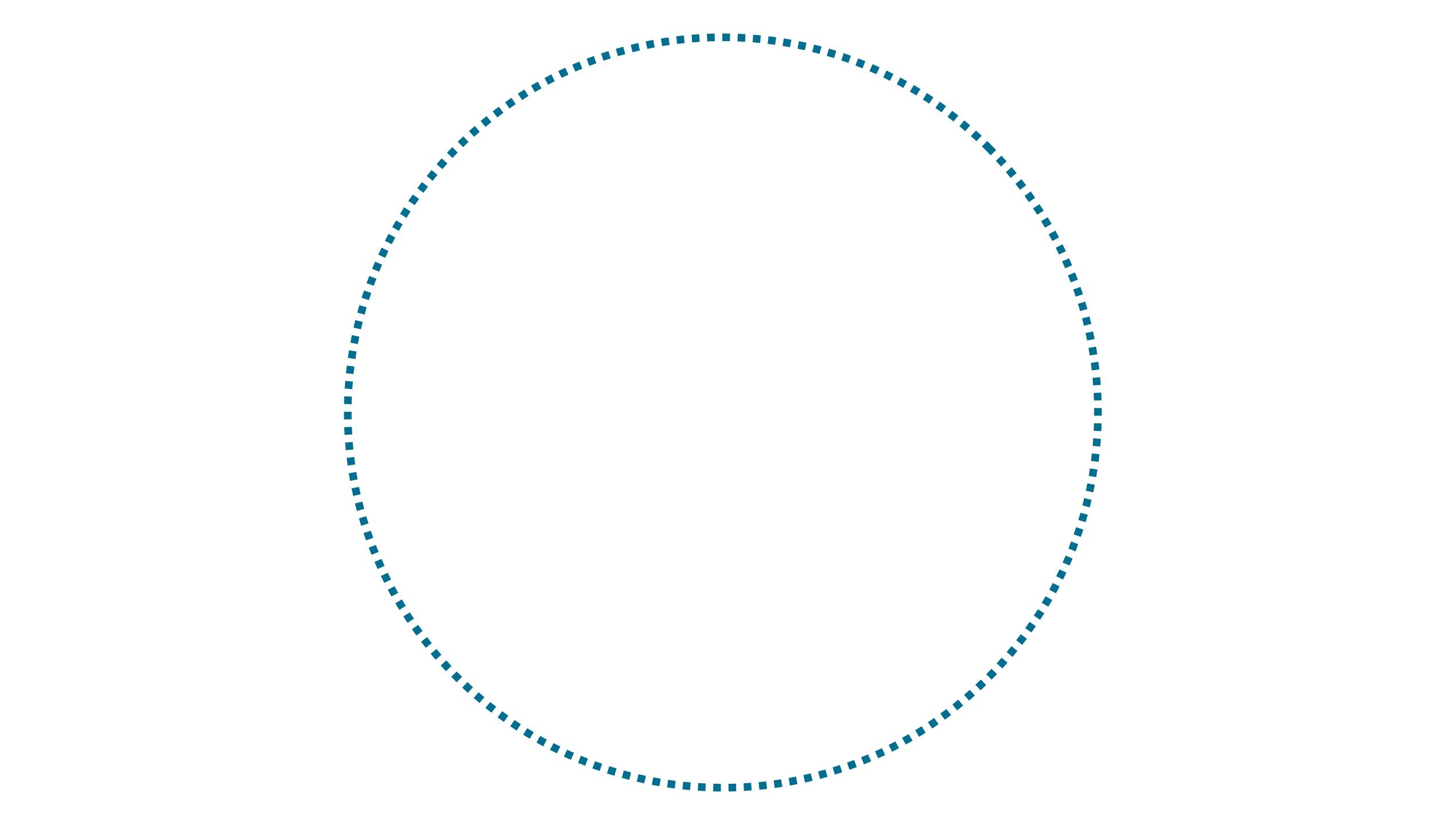




Note: Severity was assessed on a 1-7 Likert scale [1 – Low severity, 7 – High severity]

Source: World Economic Forum, Global Risks Perception Survey 2024-2025





Redefinition

Tech allows for the creation of new tasks, previously inconceivable

Modification

Tech allows for significant task redesign

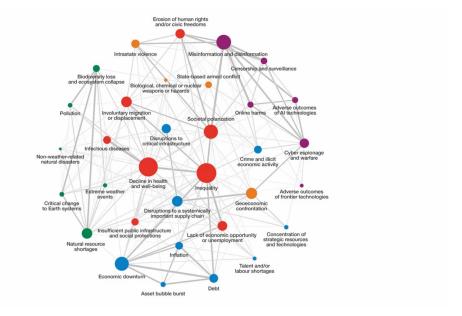
Augmentation

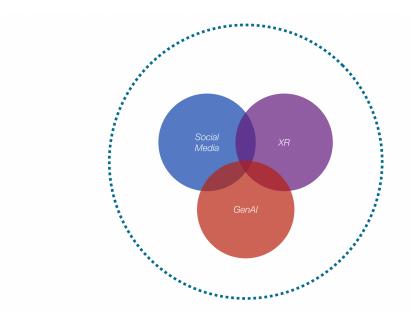
Tech acts as a direct tool substitute, with functional improvement

Substitution

Tech acts as a direct tool substitute, with no functional change

Social	Mobility	Visualization	Storytelling	Gaming
200,000 years	70,000 years	40,000 years	17,000 years	8,000 years





Analyzing Surveys – Sample Prompts

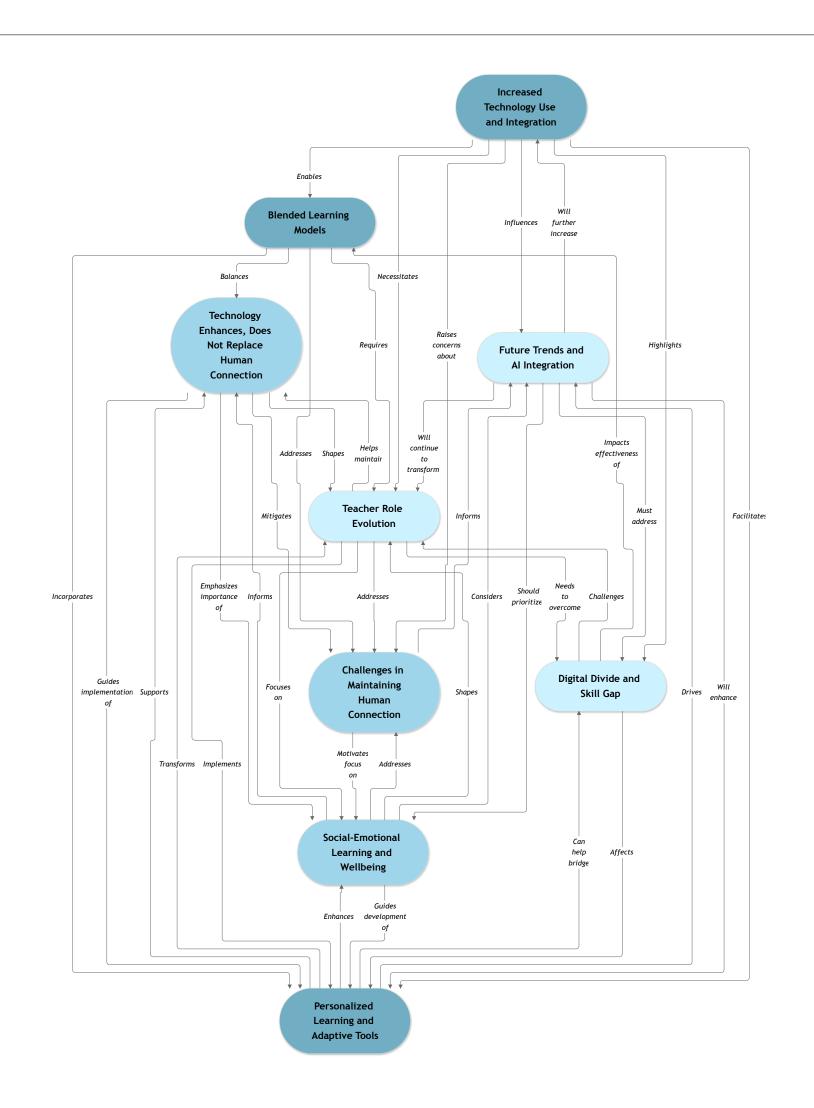
- 1. You are a sociologist researching uses of technology in education. In a survey, 82 educators were asked "In the last few years, how have you seen the balance between human connection and technology in education evolving?" Please identify the 9 most frequent themes in the responses below, grouping together similar themes under one heading, and list them, including a summary for each theme, and the total number of responses to each theme, regardless of whether multiple themes were mentioned in the same response.
- 2. Educators who were asked the question "In the last few years, how have you seen the balance between human connection and technology in education evolving?" identified the following nine themes, with their corresponding summaries:
 - <the nine themes and summaries from 1.>
 - Please create a Mermaid diagram, showing all relevant interconnections among these nine themes, making sure to label all connections with explanatory labels.
- 3. Educators who were asked the question "In the last few years, how have you seen the balance between human connection and technology in education evolving?" identified the following nine themes, with their corresponding summaries:
 - <the nine themes and summaries from 1.>
 - Please group these nine themes into three groups of three, indicating the rationale for each group.

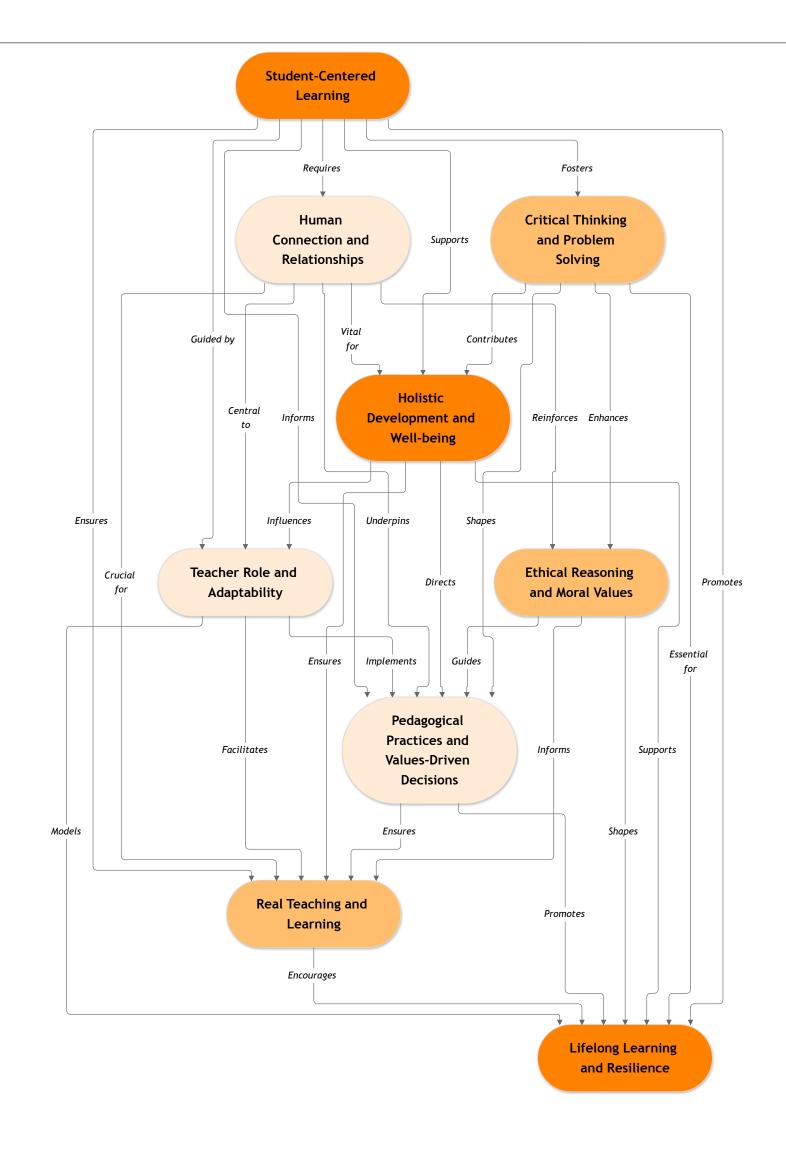
Sample Survey – Al Analysis Results

- Q1: In the last few years, how have you seen the balance between human connection and technology in education evolving?
 - Group 1: Technology Integration and Its Impact
 - Increased Technology Use and Integration
 - Blended Learning Models
 - Personalized Learning and Adaptive Tools
 - Group 2: Human Aspects and Challenges
 - Technology Enhances, Does Not Replace Human Connection
 - Challenges in Maintaining Human Connection
 - Social-Emotional Learning and Wellbeing
 - Group 3: Evolving Roles and Future Considerations
 - Teacher Role Evolution
 - Digital Divide and Skill Gap
 - Future Trends and Al Integration

- Q2: What's a core value or practice in education that you believe should remain constant, even as technology advances?
 - Group 1: Learner-Centric Approaches
 - Student-Centered Learning
 - Holistic Development and Well-being
 - · Lifelong Learning and Resilience
 - Group 2: Core Educational Values
 - Critical Thinking and Problem Solving
 - Ethical Reasoning and Moral Values
 - Real Teaching and Learning
 - Group 3: Human Elements in Education
 - Human Connection and Relationships
 - Teacher Role and Adaptability
 - Pedagogical Practices and Values-Driven Decisions

Sample Survey – Visualization





Fraction Difficulty: Bottom Number

2.5 Pro (preview)



You are a fourth grade student learning about fractions. You are asked the following question on a survey: "Please describe in a single sentence the biggest difficulty you've run into when learning about fractions." Please reply to the question.



The biggest difficulty is remembering that when the bottom number gets bigger, the piece of the whole actually gets smaller.

௴ ワ =× Ů □

Now, please create 100 responses to the question, as a different fourth grade student at the same school each time.



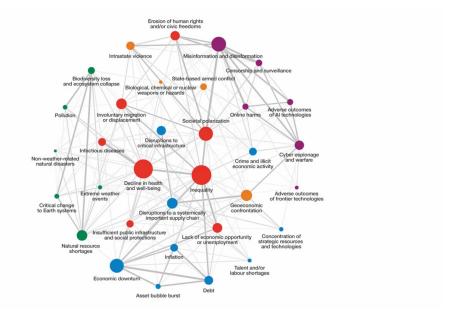
Of course, here are 100 responses to the survey question from the perspective of different fourth-grade students.

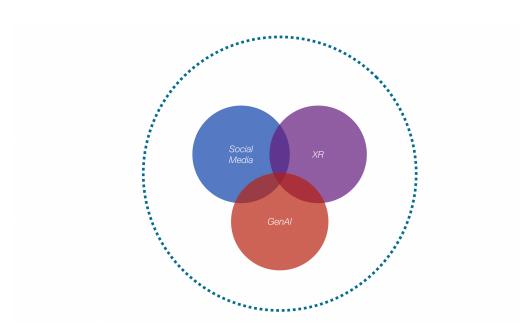
Question: "Please describe in a single sentence the biggest difficulty you've run into when learning about fractions."

Student Responses:

- 1. The hardest part is trying to get the same bottom numbers so I can add or subtract the fractions.
- 2. I don't get why you can't just add the numbers on the bottom when you're adding fractions.
- 3. It's hard to remember that a big number on the bottom means the pieces are actually smaller.
- 4. My brain gets mixed up trying to remember which number is the numerator and which is the denominator.
- 5. I never know when a fraction is simplified all the way or if I have to keep going.
- 6. Changing a mixed number like 3 $\frac{1}{2}$ into an improper fraction is really confusing.
- 7. I don't understand how 1/2 can be the same thing as 5/10 since they use different numbers.
- 8. I always forget that you're not supposed to add the bottom numbers together.
- 9. It's confusing that a fraction is one number but it has two numbers inside of it.
- 10. Comparing fractions is hard, like I can't just tell if 2/3 is bigger than 3/4 by looking.
- 11. Reading the word problems and trying to figure out what to do with the fractions is









The Illusion of Thinking: Understanding the Strengths and Limitations of Reasoning Models via the Lens of Problem Complexity

Parshin Shojaee*† Iman Mirzadeh* Keivan Alizadeh Maxwell Horton Samy Bengio Mehrdad Farajtabar

Apple

Abstract

Recent generations of frontier language models have introduced Large Reasoning Models (LRMs) that generate detailed thinking processes before providing answers. While these models demonstrate improved performance on reasoning benchmarks, their fundamental capabilities, scaling properties, and limitations remain insufficiently understood. Current evaluations primarily focus on established mathematical and coding benchmarks, emphasizing final answer accuracy. However, this evaluation paradigm often suffers from data contamination and does not provide insights into the reasoning traces' structure and quality. In this work, we systematically investigate these gaps with the help of controllable puzzle environments that allow precise manipulation of compositional complexity while maintaining consistent logical structures. This setup enables the analysis of not only final answers but also the internal reasoning traces, offering insights into how LRMs "think". Through extensive experimentation across diverse puzzles, we show that frontier LRMs face a complete accuracy collapse beyond certain complexities. Moreover, they exhibit a counterintuitive scaling limit: their reasoning effort increases with problem complexity up to a point, then declines despite having an adequate token budget. By comparing LRMs with their standard LLM counterparts under equivalent inference compute, we identify three performance regimes: (1) lowcomplexity tasks where standard models surprisingly outperform LRMs, (2) medium-complexity tasks where additional thinking in LRMs demonstrates advantage, and (3) high-complexity tasks where both models experience complete collapse. We found that LRMs have limitations in exact computation: they fail to use explicit algorithms and reason inconsistently across puzzles. We also investigate the reasoning traces in more depth, studying the patterns of explored solutions and analyzing the models' computational behavior, shedding light on their strengths, limitations, and ultimately raising crucial questions about their true reasoning capabilities.

1 Introduction

Large Language Models (LLMs) have recently evolved to include specialized variants explicitly designed for reasoning tasks—Large Reasoning Models (LRMs) such as OpenAI's o1/o3 [1, 2], DeepSeek-R1 [3], Claude 3.7 Sonnet Thinking [4], and Gemini Thinking [5]. These models are new artifacts, characterized by their "thinking" mechanisms such as long Chain-of-Thought (CoT) with self-reflection, and have demonstrated promising results across various reasoning benchmarks. Their

^{*}Equal contribution.

[†]Work done during an internship at Apple.

[{]p shojaee, imirzadeh, kalizadehvahid, mchorton, bengio, farajtabar}@apple.com



You

Please design a Reacting to the Past (RTTP) game about the early Industrial Revolution.



ChatGPT

Title: Engines of Progress: The Dawn of the Industrial Revolution

Setting:

The game is set in the late 18th and early 19th century, specifically in the period from 1760 to 1830. The location is Great Britain, the epicenter of the Industrial Revolution.

Roles:

- 1. Factory Owners: Advocates for the industrialization, their objective is to gain as much profit as possible and lobby for fewer regulations.
- 2. Workers: They are split into two groups, those who are benefiting from the changes (higher wages, more jobs) and those who are suffering (long hours, child labor, unhealthy conditions). Their objective is to improve working conditions and wages.
- 3. Luddites: Workers who are against industrialization and often resort to sabotaging machinery. Their objective is to halt the progress of industrialization and protect their traditional means of livelihood.
- 4. Politicians: Members of parliament who are divided on the issue. Some want to encourage progress and economic gravth, while others are worried about social unrest and the moral implications of d industrialization.







Ruben

Please list at least ten aspects of Flaubert's Madame Bovary that have received comparatively little attention, but are worthwhile exploring in terms of both Flaubert's world, as well as to how some of his thinking processes might be relevant to current intellectual inquiry.

Difference Thought process

Excavated nuanced, underexplored dimensions of Madame Bovary's narrative.

Underexplored Aspects of Flaubert's **Madame Bovary**

Flaubert's masterpiece "Madame Bovary" (1857) continues to yield new insights beyond the commonly analyzed themes of realism, romanticism, and feminine desire. The following aspects have received comparatively less critical attention but offer rich avenues for exploration:

1. The Medical Discourse and Healthcare **Practices**

Charles Bovary's role as a medical officer provides a window into 19thcentury healthcare. Flaubert's detailed descriptions of medical procedures (such as Charles's botched clubfoot operation) and the pharmacist Homais's pseudo-scientific pretensions reflect the tension between traditional practices and modernizing medicine. This aspect connects to contemporary discussions about medical authority, the medicalization of social problems, and the relationships between patients and healthcare providers.

2. Economic Systems and Financial **Transactions**

The intricate web of debt that eventually destroys Emma receives less attention than her romantic pursuits, yet Flaubert meticulously documents how credit systems operated in provincial France. Emma's financial decline offers insights into how capitalism was transforming rural economies. This aspect resonates with current research on debt psychology, consumer behavior, and the psychological impacts of financial precarity.

3. The Politics of Literacy and Language Acquisition

Emma's convent education and her consumption of romantic novels shape her consciousness in ways the highlight the power dynamics of literacy. The novel subtly explores we different characters' access to

Reply to Claude















Now, please list at least ten aspects of Flaubert's Bouvard et Pécuchet that have received comparatively little attention, but are worthwhile exploring in terms of both Flaubert's world, as well as to how some of his thinking processes might be relevant to current intellectual inquiry.

(1) Thought process

Excavated twelve nuanced perspectives on Flaubert's unfinished novel.

Underexplored Aspects of Flaubert's Bouvard et Pécuchet

Flaubert's final, unfinished novel "Bouvard et Pécuchet" (published posthumously in 1881) presents a rich tapestry of nineteenth-century epistemology through its two copy clerks who attempt to master various disciplines. Below are ten aspects of this work that merit further scholarly exploration, connecting Flaubert's historical context to contemporary intellectual concerns:

1. Information Management and Knowledge Organization

Flaubert's meticulous documentation of how his protagonists collect, organize, and attempt to synthesize information offers insights into nineteenth-century knowledge management. The novel's structure mirrors the challenges of organizing expanding bodies of knowledge during the industrial age. This aspect resonates with contemporary concerns about information architecture, knowledge classification systems, and the cognitive challenges of information overload in digital environments.

2. Amateur Versus Professional Expertise

The tension between amateur enthusiasm and professional specialization throughout the novel reflects the increasing institutionalization of knowledge in Flaubert's era. The protagonists' failures highlight the emerging boundaries between dilettantism and expertise. This connects to current debates about credentialism, the democratization of knowledge, and the changing nature of expertise in an era of open access information and citizen science.

3. Failure as Epistemological Method

The systematic failures that structure the novel constitute a unique epistemological method that deserves deeper analysis. Rather than simply satirizing the protagonists, Flaubert uses failure as a productive force that generates insights about knowledge acquisition. This approach anticipates contempora.

√ erspectives on productive failure

in learning the came and the cooling of arm anime antal failure in estantif a

Reply to Claude













Ruben

Now, please take the two previous lists, and create a new list of at least four items that takes components from each one of them to identify related aspects of both works that have received comparatively little attention, but are worthwhile exploring in terms of both Flaubert's world, as well as to how some of his thinking processes might be relevant to current intellectual inquiry.

Difference Thought process

Excavated nuanced literary connections across Flaubert's narrative worlds.

Interconnected Themes in Flaubert's Madame Bovary and Bouvard et Pécuchet

1. Material Culture and Knowledge Infrastructures

Both novels demonstrate Flaubert's meticulous attention to objects and their role in mediating social and intellectual life. In Madame Bovary, objects like Emma's possessions function beyond mere symbolism to reveal how material goods structured social relationships and identity formation. Similarly, in Bouvard et Pécuchet, the protagonists' books, scientific instruments, and specimens constitute a material infrastructure that shapes their cognitive processes. Together, these works offer insights into how knowledge was physically embodied in nineteenth-century France. This connection illuminates contemporary research in material culture studies, science and technology studies, and discussions about how physical and digital infrastructures continue to shape intellectual work and social status.

2. Provincial Modernity and Environmental Transformation

Flaubert's provincial settings in both novels serve as laboratories for exploring modernization's impact on rural environments and communities. Madame Bovary's Norman countryside documents agricultural transformations during France's industrialization, while Bouvard et Pécuchet's experiments with farming, landscaping, and local governance reveal tensions between urban epistemologies and rural practices. Through these parallel explorations, Flaubert documents how modernity reconfigured relationships between humans and their environments. This intersection connects to contemporary ecocritical approaches, research on spatial inequality, and questions about how different forms of knowledge mediate environmental relationships in an era of ecological crisis.

3. Embodied Experience and Physical

Reply to Claude













Now, please create for each of these four topics an outline listing the following: background knowledge a scholar should acquire in order to research it; literary works by other authors contemporary to Flaubert that are relevant to an exploration of the theme; literary works by late twentieth or early twenty-first century authors that are likewise relevant to this exploration; adaptations of Flaubert's work in other media that are relevant to this exploration. Make sure to explain why each item has been included in the outline.

Thought process

Mapped scholarly research themes for Flaubert's interconnected literary works.

Research Outlines for Interconnected Themes in Flaubert's Works

1. Material Culture and Knowledge Infrastructures

Background Knowledge Required

- History of Material Culture in 19th Century France: Essential for understanding how objects functioned socially and economically in Flaubert's context. Scholars should examine domestic interiors, fashion, and consumer goods of the period.
- 19th Century Book Production and Publishing: Necessary for appreciating how knowledge was physically manifested and distributed in Flaubert's era.
- Development of Museums and Collections: Provides context for how objects were categorized, displayed, and valued during this period of intense collecting.
- Thing Theory (Bill Brown) and Object-Oriented Ontology: Theoretical frameworks that help analyze human-object relationships beyond mere symbolism.
- **History of Scientific Instruments**: Critical for understanding the material tools that mediated scientific knowledge in Bouvard et Pécuchet's experiments.

Contemporary Literary Works

- Balzac's "La Comédie Humaine": Particularly works like "Le Père Goriot" that meticulously catalog possessions as markers of social status, providing a useful comparison to Flaubert's object descriptions.
- Zola's "Au Bonheur des Dan. : Documents the rise of

Reply to Claude



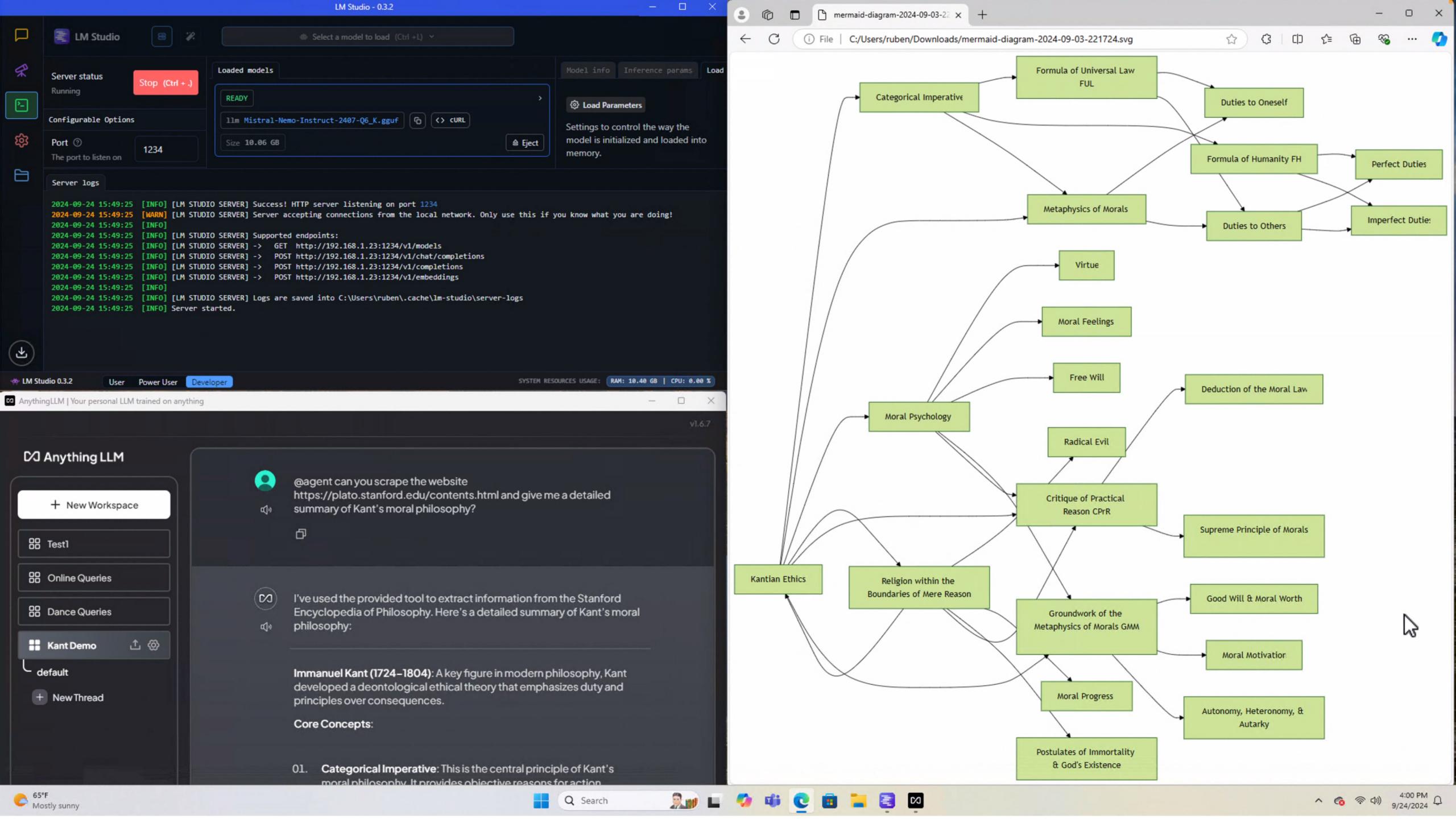














(Image Credit: Bryan Alexander/Midjourney)

Prompt: Please describe this image.

Reply (Llava Phi 3-Mini):

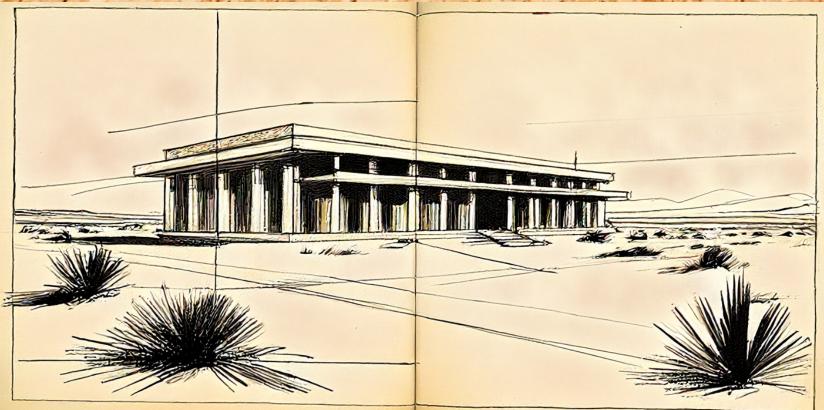
This is a black and white photograph with a dreamlike quality, as if it has been edited to emphasize its artistic nature over realism. The central figure in the photo is a man standing near an archway with columns on either side. He is looking at something off-camera to the right of the photo. Above him, there are two lights hanging from the ceiling, casting a warm glow through their reflections in a mirror that hangs above the man's head.

The floor beneath them is covered with tiles, and another set of columns can be seen on the right side of the photo, adding to the architectural symmetry of the scene. The perspective of the photo suggests it was taken from a low angle, looking up at the man standing near the archway. This creates an imposing sense of scale and grandeur in the image.

In terms of color, this is not your typical photograph as all colors have been removed, leaving only shades of gray and black. This gives the photo an ethereal quality that adds to its overall mood. The lights above are warm-toned, their reflections dancing across the mirror's surface and adding a touch of whimsy to the otherwise solemn scene.

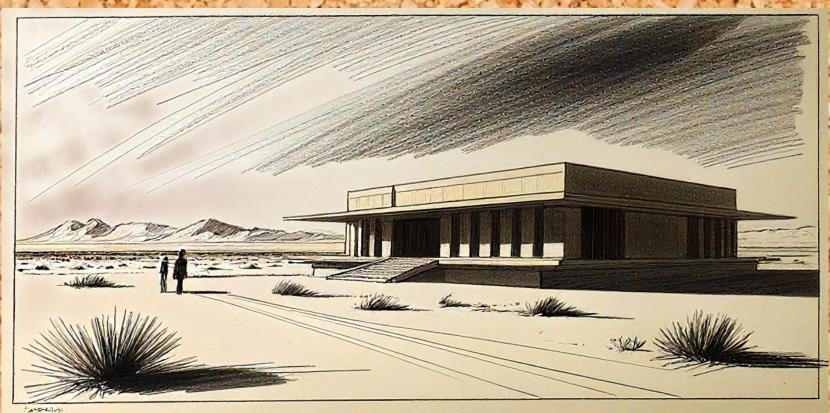
Overall, this image is a beautiful blend of architecture and artistry, capturing not just what one can see with their eyes but also inviting them to imagine what might be seen beyond sight.

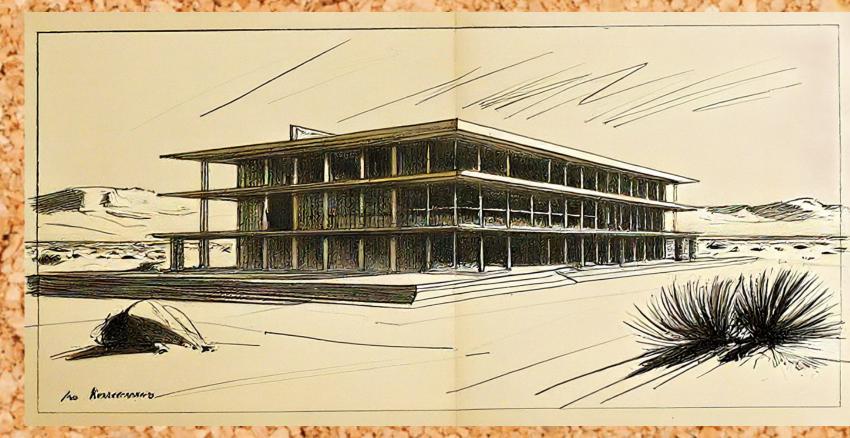




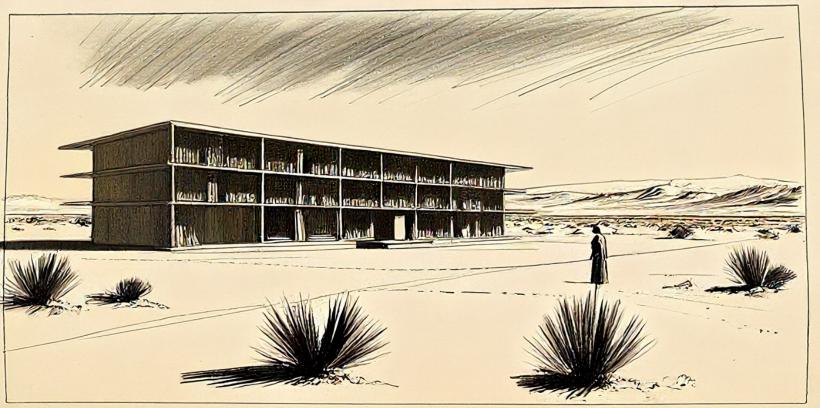






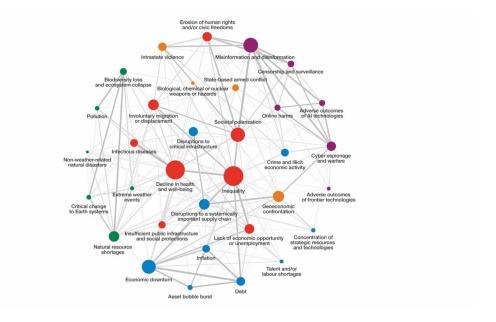


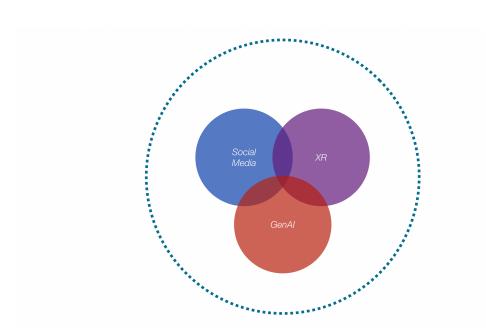






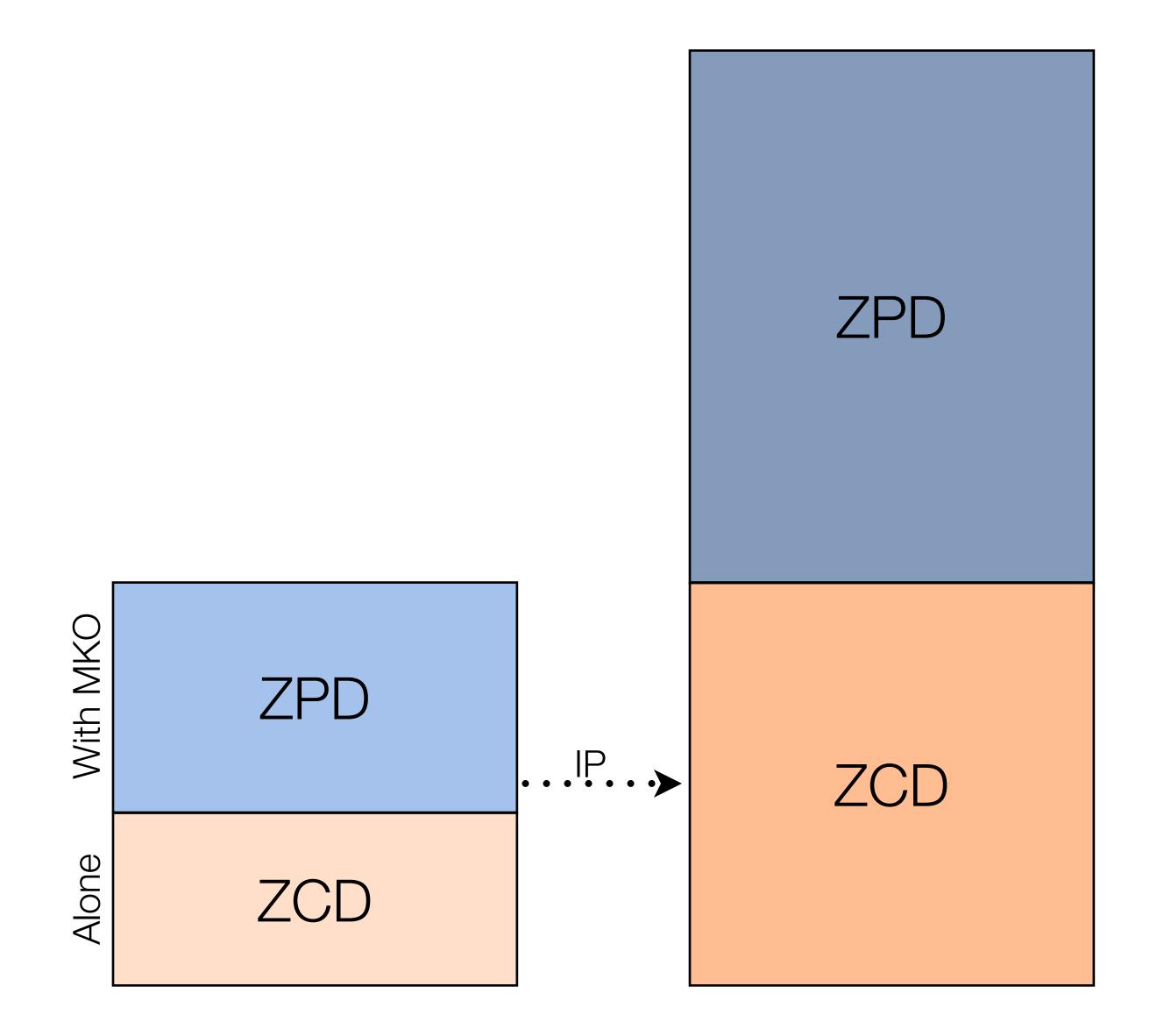


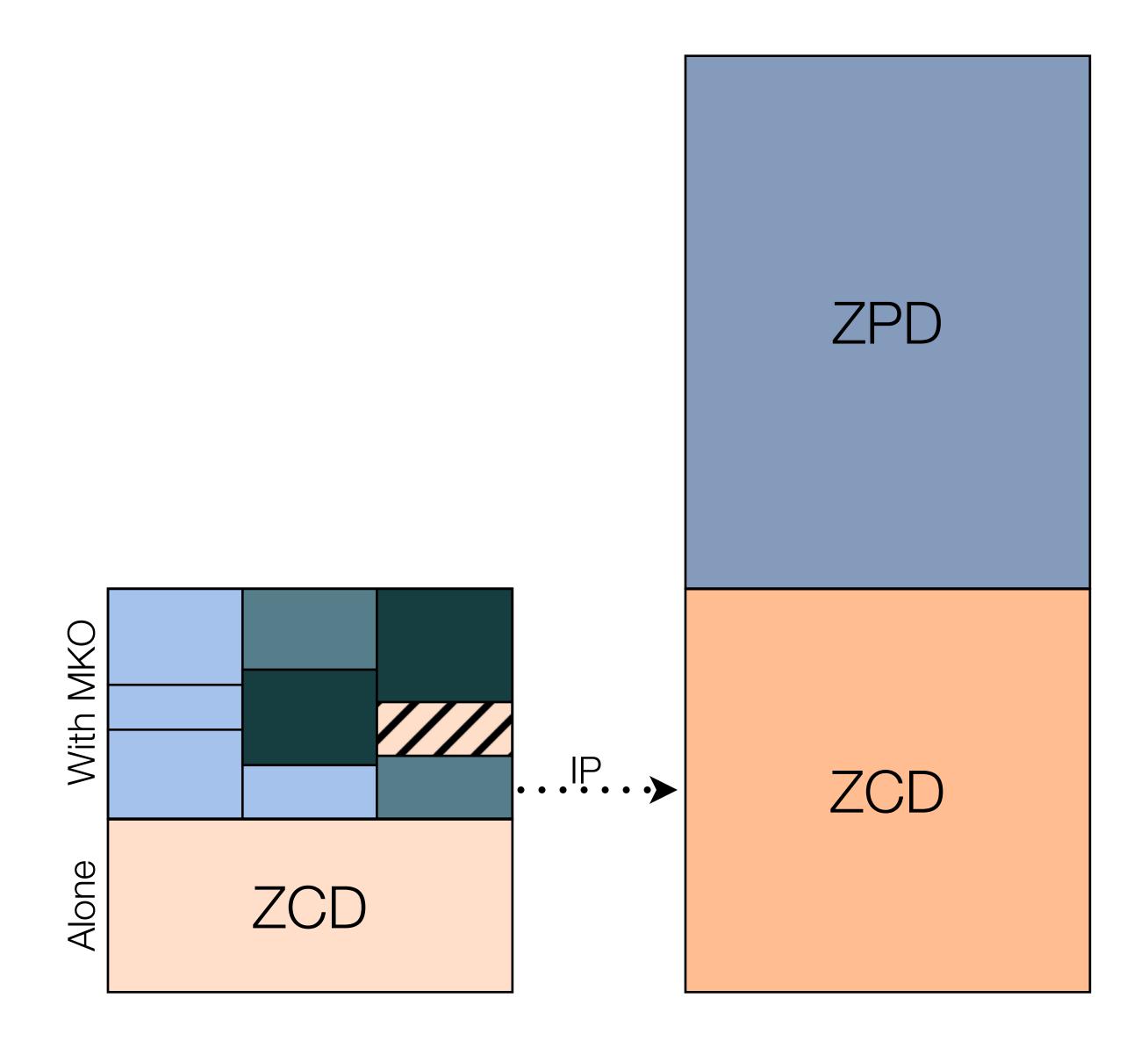




Social Learning

Practice

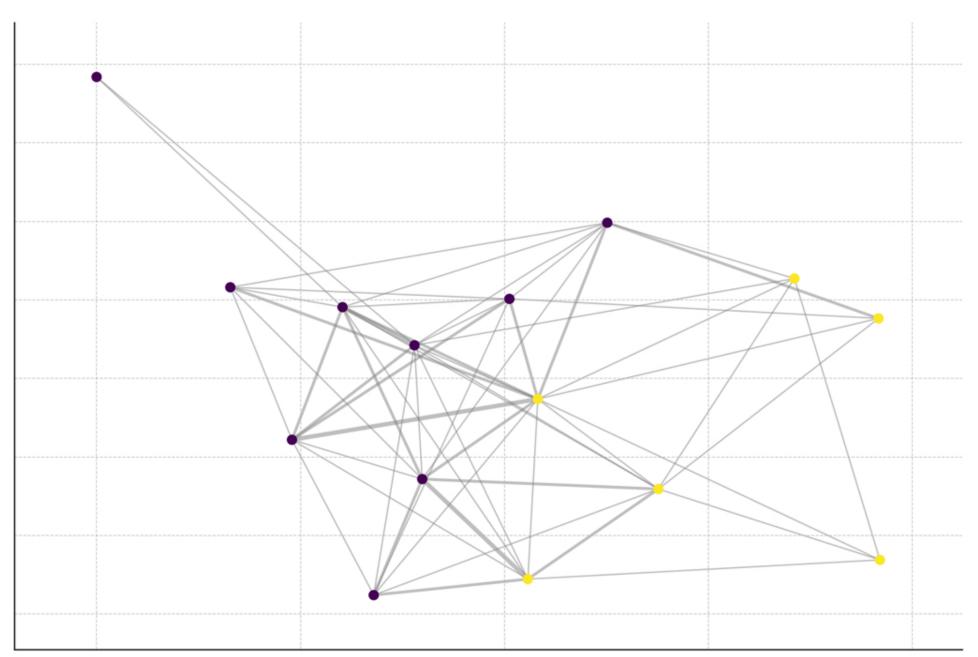






community detection. Let's proceed with that. I'll modify the code accordingly and generate the visualization. [>-]





The visualization of the weighted monopartite network has been generated. In this graph:

- Nodes are colored based on the community they belong to, detected using a Louvain-like algorithm.
- Edge thickness and color shading indicate the weights of the edges.



This visualization provides a graphical representation of the network



Message ChatGPT...



Ruben

Please write a Jupyter notebook that will provide a simple chat interface to an LLM running on a local LM Studio server. The code should be carefully commented, with all functionality clearly explained.

1) Thought process

Creating a Jupyter notebook for LLM chat interface

I'll create a Jupyter notebook that provides a simple chat interface to an LLM running on a local LM Studio server. I'll make sure the code is well-documented with clear explanations of all functionality.

LM Studio Chat Interface Notebook

Code



This Jupyter notebook creates a simple but comprehensive chat interface for interacting with a local LLM running through LM Studio. The notebook includes:

- 1. **Setup and installation** All required packages are installed and imported
- 2. Connection configuration Settings to connect to the LM Studio server (which follows OpenAI's API format)
- 3. **API interaction functions** Well-documented functions to communicate with the LLM
- 4. Conversation management A class that handles the chat history and context
- 5. User interface A clean UI built with IPython widgets including:
 - System prompt customization
 - Temperature and max tokens controls
 - Chat history display
 - Message input and send functionality
 - Reset conversation button
 - Connection status indicator

The code is thoroughly commented, explaining what each component does and how the different parts work together. Additionally, the notebook includes troubleshooting tips and suggestions for potential improvements.

To use this notebook, you'll need to ave LM Studio installed and running with a server started. The notebook will automatically connect

Reply to Claude

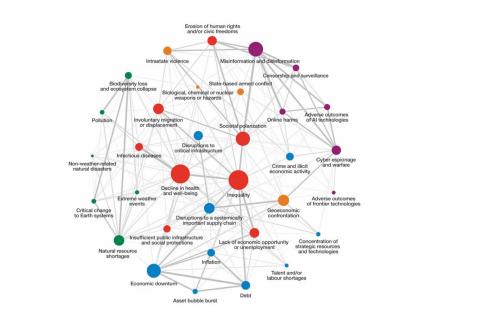


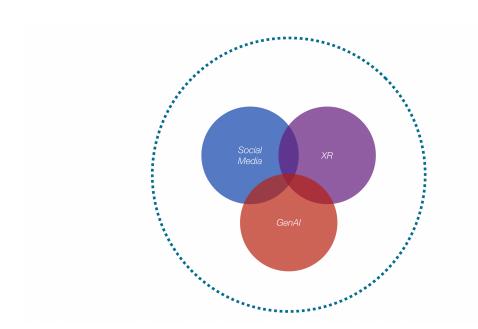












Social Learning

Practice

Metacognition

Core Thinking Routines (Harvard Project Zero)



Think, Pair, Share (Adapted)

Pose a question to students. Give students a few minutes to **think**.

Invite students to **pair** with a nearby student to **share** their thoughts.

Purpose: What kind of thinking does this routine encourage?

This routine promotes understanding through active reasoning and explanation. Because students listen to and share ideas with others, it also encourages students to understand multiple perspectives.

Application: When and where can I use it?

This routine can be used when it would help for students to process their thinking aloud with another student. For example, you may ask students to Think, Pair, Share before starting a science experiment, in the middle of solving a math problem, after reading a passage of a book, etc. Sharing can also be done in small groups.

Launch: What are some tips for starting and using this routine?

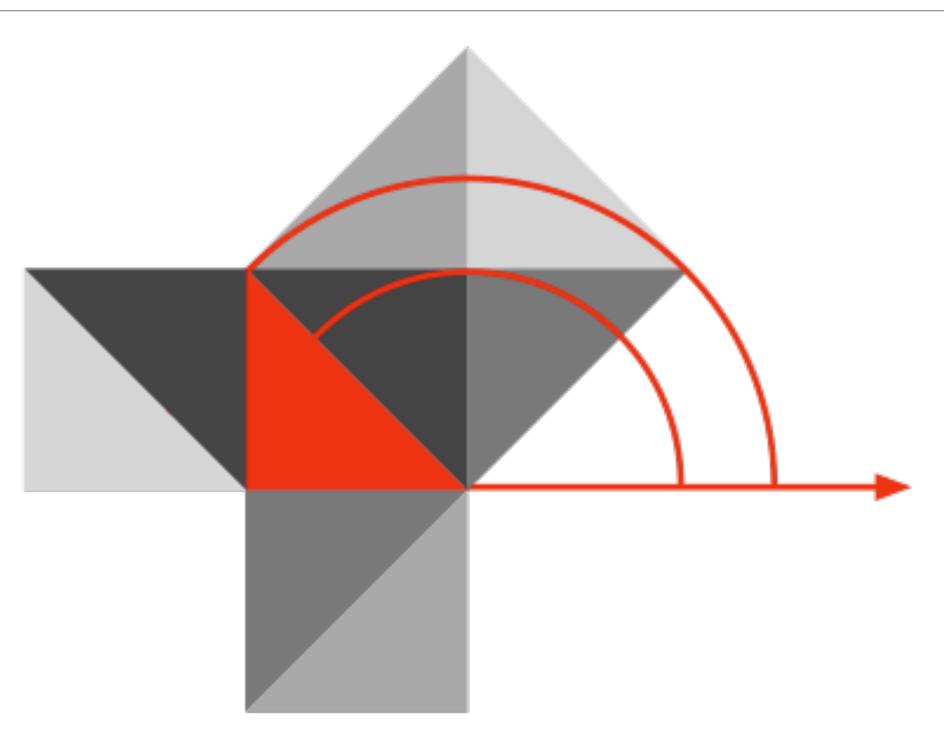
When first introducing the routine, you may want to remind students to take turns, listen carefully, and ask clarifying questions of each other. One way to encourage students to listen actively to each other is to tell them that when they have completed their conversations, you will ask some students to explain their partner's thinking. Encourage students to make their thinking visible by asking them to write or draw their ideas before and/or after sharing (perhaps in a journal). You may also want to document students' ideas and display them in the learning space. Sometimes it is useful to have pairs or groups summarize their ideas for the whole class.

Share your experience with this thinking routine on social media using the hashtags **#PZThinkingRoutines** and **#ThinkPairShare**.





Hippasus



Blog: http://hippasus.com/blog/

Email: rubenrp@hippasus.com

LinkedIn: https://www.linkedin.com/in/rubenpuentedura/

Twitter: @rubenrp

This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License.

